IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Debatin *et al.*	**Filed via EFS Web April 1, 2008**
Title:	SMAC-PEPTIDES AS THERAPEUTICS AGAINST CANCER AND AUTOIMMUNE DISEASES	
Appl. No.:	10/511,037	
§371 Date:	January 19, 2005	
I.A. Filing Date	April 17, 2003	
Examiner:	Hong Sang	
Art Unit:	1643	
Confirmation Number:	6277	

STATEMENT AND DECLARATION TO SUPPORT FILING AND SUBMISSION OF SEQUENCE LISTING BY AMENDMENT IN ACCORDANCE WITH 37 C.F.R. § 1.57(f)

Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Sir:

In connection with a Sequence Listing submitted concurrently herewith for insertion into the captioned specification, the undersigned hereby declares that the material being inserted by amendment is the material previously incorporated by reference and that the amendment contains no new matter.

The sequence listing submitted herewith includes the sequences from GenBank accession numbers disclosed in the captioned application and priority documents EP 02008199.8 and EP 02015499.3. Specifically, the sequence listing includes the Smac polypeptide of GenBank accession no. AAF87716 (SEQ ID NO:1), the polynucleotide encoding the Smac polypeptide as disclosed in GenBank accession no. AF262240 (SEQ ID NO:2), the HIV-1 tat polypeptide of GenBank accession no. CAA45921 (SEQ ID NO:3) and the polynucleotide encoding the HIV-1 Genome of GeneBank accession no. M15654 (SEQ ID NO:4).

The GenBank accession numbers referred to above may be found on page 4, line 15; page 6, lines 9-10; and page 16, line 1 of the International Publication No. WO 03/086470, which corresponds to the present application. Additionally, the GenBank accession numbers AAF87716, AF262240; and CAA45921 may be found in priority document EP 02015499.3 on page 2, line 28; page 6, lines 4-5 and page 12, lines 15-16. Finally, the GenBank accession numbers AAF87716, AF262240; and M15654 may be found in priority document EP 02008199.8 on page 2, line 28; page 6, line 5 and page 15, line 4.

Submitted herewith as Exhibits 1-4 are copies of each of the GenBank entries and their respective sequence revision histories. Genbank accession numbers AAF87716 and AF262240 have never been revised since their original GenBank submission. *See* Exhibits 1-2. GenBank accession numbers CAA45921 and M15654 have been revised, however the version numbers of the sequences have remained the same as indicated on the sequence revision history of each submission. *See* Exhibits 3 and 4. The revisions indicated in the histories for GenBank accession numbers CAA45921 and M15654 do not relate to sequence revisions. A revision of a sequence itself would have changed the version number of the sequence, according to the National Center for Biotechnology Information (NCBI). As such, the undersigned believes that the amendment to add the sequence listing filed herewith does not introduce new matter.

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent resulting therefrom.

Respectfully submitted,

Date April 1, 2008

By _____
for Michele M. Simkin
Attorney for Applicant
Registration No. 34,717
Reg # 54,161

FOLEY & LARDNER LLP
Customer Number: 22428
Telephone: (202) 672-5404
Facsimile: (202) 672-5399

EXHIBIT 1

1: AAF87716. Reports Smac [Homo sapien...[gi:9454219]

Features Sequence

```
LOCUS       AAF87716                 239 aa            linear   PRI 26-JUL-2000
DEFINITION  Smac [Homo sapiens].
ACCESSION   AAF87716
VERSION     AAF87716.1  GI:9454219
DBSOURCE    locus AF262240 accession AF262240.1
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 239)
  AUTHORS   Du,C., Fang,M., Li,Y., Li,L. and Wang,X.
  TITLE     Smac, a mitochondrial protein that promotes cytochrome c-dependent
            caspase activation by eliminating IAP inhibition
  JOURNAL   Cell 102 (1), 33-42 (2000)
   PUBMED   10929711
REFERENCE   2  (residues 1 to 239)
  AUTHORS   Du,C., Fang,M., Li,Y. and Wang,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2000) Howard Hughes Medical Institute and
            Department of Biochemstry, University of Texas Southwestern Medical
            Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA
COMMENT     Method: conceptual translation supplied by author.
FEATURES             Location/Qualifiers
     source          1..239
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     Protein         1..239
                     /product="Smac"
                     /function="binds IAPs and neutralizes their inhibition on
                     caspase activation and activity"
                     /name="antagonist of IAPs (inhibitors of apoptosis)"
     transit_peptide 1..55
                     /note="mitochondrial targeting sequence"
     Region          6..239
                     /region_name="Smac_DIABLO"
                     /note="Second Mitochondria-derived Activator of Caspases;
                     pfam09057"
                     /db_xref="CDD:72475"
     CDS             1..239
                     /coded_by="AF262240.1:20..739"
```

```
ORIGIN
        1 maalkswlsr svtsffryrq clcvpvvanf kkrcfselir pwhktvtigf gvtlcavpia
       61 qksephslss ealmrravsl vtdststfls qttyalieai teytkavytl tslyrqytsl
      121 lgkmnseeed evwqviigar aemtskhqey lklettwmta vglsemaaea ayqtgadqas
      181 itarnhiqlv klqveevhql srkaetklae aqieelrqkt qeegeeraes eqeaylred
//
```

NCBI

Sequence Revision History

Revision history for AAF87716

GI	Version	Update Date	Status
9454219	1	Jul 26 2000 12:05 AM	Live

Accession AAF87716 was first seen at NCBI on Jul 26 2000 12:05 AM

Entrez

Search for Genes
Entrez Gene provides gene-specific data for multiple taxa

Help|FAQ

Batch Entrez: Upload a file of GI or accession numbers to retrieve protein or nucleotide sequences

Check sequence revision history

How to create WWW links to Entrez

LinkOut

My NCBI (Cubby)

Related resources

BLAST

Reference sequence project

Entrez Gene

Clusters of orthologous groups

Protein reviews on the web

Disclaimer | Write to the Help Desk
NCBI | NLM|NIH

EXHIBIT 2

1: AF262240. Reports Homo sapiens Smac...[gi:9454218] Links

Features Sequence

```
LOCUS       AF262240                1358 bp    mRNA    linear   PRI 26-JUL-2000
DEFINITION  Homo sapiens Smac mRNA, complete cds; nuclear gene for
            mitochondrial product.
ACCESSION   AF262240
VERSION     AF262240.1  GI:9454218
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1358)
  AUTHORS   Du,C., Fang,M., Li,Y., Li,L. and Wang,X.
  TITLE     Smac, a mitochondrial protein that promotes cytochrome c-dependent
            caspase activation by eliminating IAP inhibition
  JOURNAL   Cell 102 (1), 33-42 (2000)
   PUBMED   10929711
REFERENCE   2  (bases 1 to 1358)
  AUTHORS   Du,C., Fang,M., Li,Y. and Wang,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2000) Howard Hughes Medical Institute and
            Department of Biochemstry, University of Texas Southwestern Medical
            Center, 5323 Harry Hines Blvd., Dallas, TX 75235, USA
FEATURES             Location/Qualifiers
     source          1..1358
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             20..739
                     /function="binds IAPs and neutralizes their inhibition on
                     caspase activation and activity"
                     /note="antagonist of IAPs (inhibitors of apoptosis)"
                     /codon_start=1
                     /product="Smac"
                     /protein_id="AAF87716.1"
                     /db_xref="GI:9454219"
                     /translation="MAALKSWLSRSVTSFFRYRQCLCVPVVANFKKRCFSELIRPWHK
                     TVTIGFGVTLCAVPIAQKSEPHSLSSEALMRRAVSLVTDSTSTFLSQTTYALIEAITE
                     YTKAVYTLTSLYRQYTSLLGKMNSEEEDEVWQVIIGARAEMTSKHQEYLKLETTWMTA
                     VGLSEMAAEAAYQTGADQASITARNHIQLVKLQVEEVHQLSRKAETKLAEAQIEELRQ
                     KTQEEGEERAESEQEAYLRED"
     transit_peptide 20..184
                     /note="mitochondrial targeting sequence"
```

```
ORIGIN
        1 ggcgtccgcg cgctgcacaa tggcggctct gaagagttgg ctgtcgcgca gcgtaacttc
       61 attcttcagg tacagacagt gtttgtgtgt tcctgttgtg gctaacttta agaagcggtg
      121 tttctcagaa ttgataagac catggcacaa aactgtgacg attggctttg gagtaaccct
      181 gtgtgcggtt cctattgcac agaaatcaga gcctcattcc cttagtagtg aagcattgat
      241 gaggagagca gtgtctttgg taacagatag cacctctacc tttctctctc agaccacata
      301 tgcgttgatt gaagctatta ctgaatatac taaggctgtt tataccttaa cttctcttta
      361 ccgacaatat acaagtttac ttgggaaaat gaattcagag gaggaagatg aagtgtggca
      421 ggtgatcata ggagccagag ctgagatgac ttcaaaacac caagagtact tgaagctgga
      481 aaccacttgg atgactgcag ttggtctttc agagatggca gcagaagctg catatcaaac
      541 tggcgcagat caggcctcta taaccgccag gaatcacatt cagctggtga aactgcaggt
      601 ggaagaggtg caccagctct cccggaaagc agaaaccaag ctggcagaag cacagataga
      661 agagctccgt cagaaaacac aggaggaagg ggaggagcgg gctgagtcgg agcaggaggc
      721 ctacctgcgt gaggattgag ggcctgagca cactgccctg tctccccact cagtggggaa
      781 agcaggggca gatgccaccc tgcccagggt tggcatgact gtctgtgcac cgagaagagg
      841 cggcaggtcc tgccctggcc aatcaggcga gacgcctttg tgagctgtga gtgcctcctg
      901 tggtctcagg cttgcgctgg acctggttct tagcccttgg gcactgcacc ctgtttaaca
      961 tttcacccca ctctgtacag ctgctcttac ccatttttt tacctcacac ccaaagcatt
     1021 ttgcctacct gggtcagaga gaggagtcct ttttgtcatg cccttaagtt cagcaactgt
     1081 ttaacctgtt ttcagtctta tttacgtcgt caaaaatgat ttagtacttg ttcccctctgt
     1141 tgggatgcca gttgtggcag ggggagggga acctgtccag tttgtacgat ttctttgtat
     1201 gtatttctga tgtgttctct gatctgcccc cactgtcctg tgaggacagc tgaggccaag
     1261 gagtgaaaaa cctattacta ctaagagaag gggtgcagag tgtttacctg gtgctctcaa
     1321 caggacttaa catcaacagg acttaacaca gaaaaaaa
//
```

NCBI

Sequence Revision History

Revision history for AF262240

GI	Version	Update Date	Status
9454218	1	Jul 26 2000 12:05 AM	Live

Accession AF262240 was first seen at NCBI on Jul 26 2000 12:05 AM

Entrez

Search for Genes
Entrez Gene provides gene-specific data for multiple taxa

Help|FAQ

Batch Entrez: Upload a file of GI or accession numbers to retrieve protein or nucleotide sequences

Check sequence revision history

How to create WWW links to Entrez

LinkOut

My NCBI (Cubby)

Related resources

BLAST

Reference sequence project

Entrez Gene

Clusters of orthologous groups

Protein reviews on the web

Disclaimer | Write to the Help Desk
NCBI | NLM|NIH

EXHIBIT 3

1: CAA45921. Reports HIV-1 tat [Human ...[gi:60145]

BLink, Conserved Domains, Links

Features Sequence

```
LOCUS       CAA45921                 86 aa            linear   VRL 20-NOV-1992
DEFINITION  HIV-1 tat [Human immunodeficiency virus 1].
ACCESSION   CAA45921
VERSION     CAA45921.1  GI:60145
DBSOURCE    embl accession X64650.1
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retro-transcribing viruses; Retroviridae;
            Orthoretrovirinae; Lentivirus; Primate lentivirus group.
REFERENCE   1  (residues 1 to 86)
  AUTHORS   Siderovski,D.P., Matsuyama,T., Frigerio,E., Chui,S., Min,X.,
            Erfle,H., Sumner-Smith,M., Barnett,R.W. and Mak,T.W.
  TITLE     Random mutagenesis of the human immunodeficiency virus type-1
            trans-activator of transcription (HIV-1 Tat)
  JOURNAL   Nucleic Acids Res. 20 (20), 5311-5320 (1992)
   PUBMED   1437550
REFERENCE   2  (residues 1 to 86)
  AUTHORS   Siderovski,D.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1992) D.P. Siderovski, Dept of Medical
            Biophysics, University of Toronto, 500 Sherbourne Street, Toronto
            Ontario M4X 1K9, CANADA
FEATURES             Location/Qualifiers
     source          1..86
                     /organism="Human immunodeficiency virus 1"
                     /db_xref="taxon:11676"
     Protein         1..86
                     /product="HIV-1 tat"
     Region          2..>47
                     /region_name="Tat"
                     /note="Transactivating regulatory protein (Tat). The
                     retroviral Tat protein binds to the Tar RNA. This
                     activates transcriptional initiation and elongation from
                     the LTR promoter. Binding is mediated by an arginine rich
                     region; pfam00539"
                     /db_xref="CDD:84847"
     CDS             1..86
                     /gene="HIV-1 tat"
                     /coded_by="X64650.1:39..299"
                     /db_xref="InterPro:IPR001831"
ORIGIN
```

```
        1 mepvdprlep wkhpgsqpkt actncyckkc cfhcqvcfit kalgisygrk krrqrrrppq
       61 gsqthqvsls kqptsqsrgd ptgpke
//
```

NCBI

Sequence Revision History

Revision history for CAA45921

GI	Version	Update Date	Status	I	II
60145	1	Oct 20 2006 8:41 AM	Live	◉	○
60145	1	Aug 8 2003 11:42 PM	Dead	○	◉
60145	1	Oct 19 2002 2:32 PM	Dead	○	○
60145	1	Mar 8 1999 10:26 PM	Dead	○	○
60145	1	May 25 1995 12:46 PM	Dead	○	○
60145	1	Dec 1 1994 11:51 AM	Dead	○	○
60145	1	Sep 1 1993 8:12 PM	Dead	○	○
60145	1	Apr 21 1993 5:17 PM	Dead	○	○

Accession CAA45921 was first seen at NCBI on Apr 21 1993 5:17 PM

EXHIBIT 4

1: M15654. Reports Human immunodefic...[gi:326383] Links

Comment Features Sequence

```
LOCUS       HIVBH102                8932 bp ss-RNA     linear   VRL 02-AUG-2006
DEFINITION  Human immunodeficiency virus type 1, isolate BH10, genome.
ACCESSION   M15654 K02008 K02009 K02010
VERSION     M15654.1  GI:326383
KEYWORDS    TAR region; acquired immune deficiency syndrome; env protein; gag
            protein; long terminal repeat (LTR); pol protein; polyprotein;
            proviral gene; reverse transcriptase; transactivator.
SEGMENT     2 of 2
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retro-transcribing viruses; Retroviridae;
            Orthoretrovirinae; Lentivirus; Primate lentivirus group.
REFERENCE   1  (bases 1 to 8932)
  AUTHORS   Wong-Staal,F., Gallo,R.C., Chang,N.T., Ghrayeb,J., Papas,T.S.,
            Lautenberger,J.A., Pearson,M.L., Petteway,S.R.Jr., Ivanoff,L.,
            Baumeister,K., Whitehorn,E.A., Rafalski,J.A., Doran,E.R.,
            Josephs,S.J., Starcich,B., Livak,K.J., Patarca,R., Haseltine,W.A.
            and Ratner,L.
  TITLE     Complete nucleotide sequence of the AIDS virus, HTLV-III
  JOURNAL   Nature 313 (6000), 277-284 (1985)
   PUBMED   2578615
REFERENCE   2  (bases 1 to 8932)
  AUTHORS   van Beveren,C.P., Coffin,J. and Hughes,S.
  TITLE     Appendix B: HTLV-3 genome
  JOURNAL   (in) Weiss,R.L., Teich,N., Varmus,H. and Coffin,J. (Eds.);
            RNA TUMOR VIRUSES, MOLECULAR BIOLOGY OF TUMOR VIRUSES, SECOND
            EDITION, 2;
            Cold Spring Harbor Laboratory, CSH, NY (1985)
REFERENCE   3  (sites)
  AUTHORS   Hostomsky,Z., Hostomska,Z., Hudson,G.O., Moomaw,E.W. and Nodes,B.R.
  TITLE     Reconstitution in vitro of RNase H activity by using purified
            N-terminal and C-terminal domains of human immunodeficiency virus
            type 1 reverse transcriptase
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 88 (4), 1148-1152 (1991)
   PUBMED   1705027
COMMENT     Original source text: Human immunodeficiency virus type 1 (HIV-1)
            proviral DNA clone BH10.
            [(in) Weiss,R., Teich,N., Varmus,H. and Coffin,J. (Eds.);RNA Tumor
            Viruses,Molecu]  review; bases 1 to 8932.
            The BH10 sequence differs from BH8 and BH5 by 0.9% in the coding
            regions and 1.8% in the noncoding regions, and the authors of [1]
            believe that these are stable variants.
            The HTLV-III genome encodes at least seven proteins: gag, pol, env,
```

tat, trs, 27K antigen and the sor 23K product. The 3' ORF (positions 8153-8773) is truncated in BH10 (stop codon at positions 8522-8524), but reads through in BH8 and other sequences to yield what is now called the 27K antigen.
The mechanism for pol gene translation has not been elucidated: a gag-pol fusion protein is possible; splicing or frameshift have not been ruled out. The viral protease would be determined by the region in question.
The Tat protein (trans-activator protein, approximately 14 kb) is an effector of an autostimulatory pathway through interaction with a positive control element, the trans-activating responsive sequence, TAR. Tat seems to be a transcriptional control molecule in HTLV-I, but is both that and a post-transcriptional regulatory molecule in HTLV-III. Deletion mutants in the tat gene are incapable of prolific replication and exhibit no cytopathic effects in T4+ cell lines.
In addition to the
9.4 kb genomic mRNA, subgenomic mRNAs of 7.4, 5.5, 5.0, 4.3, 2.0 and 1.8 kb have been detected.

FEATURES Location/Qualifiers
source 1..8932
/organism="Human immunodeficiency virus 1"
/mol_type="genomic RNA"
/db_xref="taxon:11676"
prim_transcript <1..>8932
/note="genomic mRNA"
prim_transcript <1..>8932
/note="tat, rev, nef subgenomic mRNA"
intron 66..5135
/note="tat, rev, nef subgenomic mRNA intron 1"
gene 112..1650
/gene="gag"
CDS 112..1650
/gene="gag"
/codon_start=1
/product="gag polyprotein precursor"
/protein_id="AAA44201.1"
/db_xref="GI:326388"
/translation="MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERF
AVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALD
KIEEEQNKSKKKAQQAAADTGHSSQVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVE
EKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPV
HAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTIL
KALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNTATIMMQRGNFRNQRKMVKC
FNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSYKGRPGNFLQ
SRPEPTAPPFLQSRPEPTAPPEESFRSGVETTTPPQKQEPIDKELYPLTSLRSLFGND
PSSQ"
gene <1407..4454
/gene="pol"
CDS <1407..4454
/gene="pol"
/note="ORF for pol polyprotein containing stop codons"
/codon_start=1
/product="pol polyprotein"
/protein_id="AAA44198.1"
/db_xref="GI:326385"
/translation="FFREDLAFLQGKAREFSSEQTRANSPTISSEQTRANSPTRRELQ
VWGRDNNSPSEAGADRQGTVSFNFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLE
EMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQ

```
                     IGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIG
                     PENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVL
                     DVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKI
                     LEPFKKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEP
                     PFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLL
                     RGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQI
                     YQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWE
                     TWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGY
                     VTNKGRQKVVPLTNTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDKSES
                     ELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKILFLDGIDKAQDEHE
                     KYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGK
                     VILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTSATVKAA
                     CWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKR
                     KGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKG
                     EGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED"
gene                 4399..4977
                     /gene="vif"
CDS                  4399..4977
                     /gene="vif"
                     /codon_start=1
                     /product="vif protein"
                     /protein_id="AAA44202.1"
                     /db_xref="GI:326389"
                     /translation="MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHY
                     ESPHPRISSEVHIPLGDARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPEL
                     ADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAGHNKVGSLQYLALAALITPKKIK
                     PPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH"
gene                 4917..5153
                     /gene="vpr"
CDS                  4917..5153
                     /gene="vpr"
                     /codon_start=1
                     /product="vpr protein"
                     /protein_id="AAA44203.1"
                     /db_xref="GI:326390"
                     /translation="MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWLHGLGQ
                     HIYETYGDTWAGVEAIIRILQQLLFIHFQNWVST"
gene                 5189..7779
                     /gene="tat"
CDS                  join(5189..5403,7734..7779)
                     /gene="tat"
                     /codon_start=1
                     /product="tat protein"
                     /protein_id="AAA44199.1"
                     /db_xref="GI:326386"
                     /translation="MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALG
                     ISYGRKKRRQRRRPPQGSQTHQVSLSKQPTSQSRGDPTGPKE"
gene                 5328..8008
                     /gene="rev"
CDS                  join(5328..5403,7734..8008)
                     /gene="rev"
                     /codon_start=1
                     /product="rev protein"
                     /protein_id="AAA44200.1"
                     /db_xref="GI:326387"
                     /translation="MAGRSGDSDEDLLKAVRLIKFLYQSNPPPNPEGTRQARRNRRRR
                     WRERQRQIHSISERILSTYLGRSAEPVPLQLPPLERLTLDCNEDCGTSGTQGVGSPQI
                     LVESPTVLESGAKE"
intron               5404..7733
                     /note="tat, rev, nef subgenomic mRNA intron 2"
```

```
     gene            5420..5665
                     /gene="vpu"
     CDS             5420..5665
                     /gene="vpu"
                     /codon_start=1
                     /product="vpu protein"
                     /protein_id="AAA44204.1"
                     /db_xref="GI:326391"
                     /translation="MQPIQIAIVALVVAIIIAIVVWSIVIIEYRKILRQRKIDRLIDR
                     LIERAEDSGNESEGEISALVEMGVEMGHHAPWDVDDL"
     gene            5580..8150
                     /gene="env"
     CDS             5580..8150
                     /gene="env"
                     /codon_start=1
                     /product="envelope polyprotein"
                     /protein_id="AAA44205.1"
                     /db_xref="GI:326392"
                     /translation="MRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPV
                     WKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDMVE
                     QMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKGEIKNCSFN
                     ISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYTLTSCNTSVITQACPKVSFEPIPIHYC
                     APAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSAN
                     FTDNAKTIIVQLNQSVEINCTRPNNNTRKSIRIQRGPGRAFVTIGKIGNMRQAHCNIS
                     RAKWNNTLKQIDSKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFN
                     STWFNSTWSTKGSNNTEGSDTITLPCRIKQIINMWQEVGKAMYAPPISGQIRCSSNIT
                     GLLLTRDGGNSNNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQR
                     EKRAVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLL
                     QLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWN
                     NMTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYI
                     KLFIMIVGGLVGLRIVFAVLSVVNRVRQGYSPLSFQTHLPIPRGPDRPEGIEEEGGER
                     DRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWN
                     LLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQGAYRAIRHIPRRIRQGLERILL"
     gene            8152..8523
                     /gene="nef"
     CDS             8152..8523
                     /gene="nef"
                     /note="first expressed exon; premature termination"
                     /number=3
                     /codon_start=1
                     /product="nef protein"
                     /protein_id="AAA44206.1"
                     /db_xref="GI:326393"
                     /translation="MGGKWSKSSVVGWPAVRERMRRAEPAADGVGAASRDLEKHGAIT
                     SSNTAANNADCAWLEAQEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIH
                     SQRRQDILDLWIYHTQGYFPD"
     LTR             8441..>8932
                     /note="3' LTR"
     repeat_region   8894..>8932
                     /note="R repeat 3' copy"
ORIGIN      About 120 bp downstream from HIVBH101.
        1 agctctctcg acgcaggact cggcttgctg aagcgcgcac ggcaagaggc gaggggcggc
       61 gactggtgag tacgccaaaa attttgacta gcggaggcta gaaggagaga gatgggtgcg
      121 agagcgtcag tattaagcgg gggagaatta gatcgatggg aaaaaattcg gttaaggcca
      181 gggggaaaga aaaaatataa attaaaacat atagtatggg caagcaggga gctagaacga
      241 ttcgcagtta atcctggcct gttagaaaca tcagaaggct gtagacaaat actgggacag
      301 ctacaaccat cccttcagac aggatcagaa gaacttagat cattatataa tacagtagca
      361 accctctatt gtgtgcatca aaggatagag ataaaagaca ccaaggaagc tttagacaag
      421 atagaggaag agcaaaacaa aagtaagaaa aaagcacagc aagcagcagc tgacacagga
      481 cacagcagtc aggtcagcca aaattaccct atagtgcaga acatccaggg gcaaatggta
```

```
 541 catcaggcca tatcacctag aactttaaat gcatgggtaa aagtagtaga agagaaggct
 601 ttcagcccag aagtaatacc catgttttca gcattatcag aaggagccac cccacaagat
 661 ttaaacacca tgctaaacac agtgggggga catcaagcag ccatgcaaat gttaaaagag
 721 accatcaatg aggaagctgc agaatgggat agagtacatc cagtgcatgc agggcctatt
 781 gcaccaggcc agatgagaga accaagggga agtgacatag caggaactac tagtaccctt
 841 caggaacaaa taggatggat gacaaataat ccacctatcc cagtaggaga aatttataaa
 901 agatggataa tcctgggatt aaataaaata gtaagaatgt atagccctac cagcattctg
 961 gacataagac aaggaccaaa agaacctttt agagactatg tagaccggtt ctataaaact
1021 ctaagagccg agcaagcttc acaggaggta aaaaattgga tgacagaaac cttgttggtc
1081 caaaatgcga acccagattg taagactatt ttaaaagcat tgggaccagc ggctacacta
1141 gaagaaatga tgacagcatg tcagggagta ggaggacccg gccataaggc aagagttttg
1201 gctgaagcaa tgagccaagt aacaaataca gctaccataa tgatgcagag aggcaatttt
1261 aggaaccaaa gaaagatggt taagtgtttc aattgtggca aagaagggca cacagccaga
1321 aattgcaggg cccctaggaa aaagggctgt tggaaatgtg gaaaggaagg acaccaaatg
1381 aaagattgta ctgagagaca ggctaatttt ttagggaaga tctggccttc ctacaaggga
1441 aggccaggga attttcttca gagcagacca gagccaacag ccccaccatt tcttcagagc
1501 agaccagagc caacagcccc accagaagag agcttcaggt ctggggtaga gacaacaact
1561 ccccctcaga agcaggagcc gatagacaag gaactgtatc ctttaacttc cctcagatca
1621 ctctttggca acgacccctc gtcacaataa agataggggg gcaactaaag gaagctctat
1681 tagatacagg agcagatgat acagtattag aagaaatgag tttgccagga agatggaaac
1741 caaaaatgat agggggaatt ggaggtttta tcaaagtaag acagtatgat cagatactca
1801 tagaaatctg tggacataaa gctataggta cagtattagt aggacctaca cctgtcaaca
1861 taattggaag aaatctgttg actcagattg gttgcacttt aaattttccc attagcccta
1921 ttgagactgt accagtaaaa ttaaagccag gaatggatgg cccaaaagtt aaacaatggc
1981 cattgacaga agaaaaaata aaagcattag tagaaatttg tacagaaatg gaaaaggaag
2041 ggaaaatttc aaaaattggg cctgagaatc catacaatac tccagtattt gccataaaga
2101 aaaaagacag tactaaatgg agaaaattag tagatttcag agaacttaat aagagaactc
2161 aagacttctg ggaagttcaa ttaggaatac cacatcccgc agggttaaaa aagaaaaaat
2221 cagtaacagt actggatgtg ggtgatgcat atttttcagt tcccttagat gaagacttca
2281 ggaagtatac tgcatttacc atacctagta taaacaatga gacaccaggg attagatatc
2341 agtacaatgt gcttccacag ggatggaaag gatcaccagc aatattccaa agtagcatga
2401 caaaaatctt agagcctttt aaaaaacaaa atccagacat agttatctat caatacatgg
2461 atgatttgta tgtaggatct gacttagaaa tagggcagca tagaacaaaa atagaggagc
2521 tgagacaaca tctgttgagg tggggactta ccacaccaga caaaaaacat cagaaagaac
2581 ctccattcct ttggatgggt tatgaactcc atcctgataa atggacagta cagcctatag
2641 tgctgccaga aaaagacagc tggactgtca atgacataca gaagttagtg gggaaattga
2701 attgggcaag tcagatttac ccagggatta aagtaaggca attatgtaaa ctccttagag
2761 gaaccaaagc actaacagaa gtaataccac taacagaaga agcagagcta gaactggcag
2821 aaaacagaga gattctaaaa gaaccagtac atggagtgta ttatgaccca tcaaaagact
2881 taatagcaga aatacagaag caggggcaag gccaatggac atatcaaatt tatcaagagc
2941 catttaaaaa tctgaaaaca ggaaaatatg caagaatgag gggtgcccac actaatgatg
3001 taaaacaatt aacagaggca gtgcaaaaaa taaccacaga aagcatagta atatggggaa
3061 agactcctaa atttaaacta cccatacaaa aggaaacatg ggaaacatgg tggacagagt
3121 attggcaagc cacctggatt cctgagtggg agtttgttaa tacccctcct ttagtgaaat
3181 tatggtacca gttagagaaa gaacccatag taggagcaga aaccttctat gtagatgggg
3241 cagctaacag ggagactaaa ttaggaaaag caggatatgt tactaacaaa ggaagacaaa
3301 aggttgtccc cctaactaac acaacaaatc agaaaactga gttacaagca atttatctag
3361 ctttgcagga ttcaggatta gaagtaaaca tagtaacaga ctcacaatat gcattaggaa
3421 tcattcaagc acaaccagat aaaagtgaat cagagttagt caatcaaata atagagcagt
3481 taataaaaaa ggaaaaggtc tatctggcat gggtaccagc acacaaagga attggaggaa
3541 atgaacaagt agataaatta gtcagtgctg gaatcaggaa aatactattt ttagatggaa
3601 tagataaggc ccaagatgaa catgagaaat atcacagtaa ttggagagca atggctagtg
3661 attttaacct gccacctgta gtagcaaaag aaatagtagc cagctgtgat aaatgtcagc
3721 taaaaggaga agccatgcat ggacaagtag actgtagtcc aggaatatgg caactagatt
3781 gtacacattt agaaggaaaa gttatcctgg tagcagttca tgtagccagt ggatatatag
3841 aagcagaagt tattccagca gaaacagggc aggaaacagc atattttctt ttaaaattag
3901 caggaagatg gccagtaaaa acaatacata cagacaatgg cagcaatttc accagtgcta
3961 cggttaaggc cgcctgttgg tgggcgggaa tcaagcagga atttggaatt ccctacaatc
4021 cccaaagtca aggagtagta gaatctatga ataaagaatt aaagaaaatt ataggacagg
4081 taagagatca ggctgaacat cttaagacag cagtacaaat ggcagtattc atccacaatt
```

```
4141 ttaaaagaaa agggggggatt ggggggtaca gtgcagggga aagaatagta gacataatag
4201 caacagacat acaaactaaa gaattacaaa aacaaattac aaaaattcaa aattttcggg
4261 tttattacag ggacagcaga aatccacttt ggaaaggacc agcaaagctc ctctggaaag
4321 gtgaaggggc agtagtaata caagataata gtgacataaa agtagtgcca agaagaaaag
4381 caaagatcat tagggattat ggaaaacaga tggcaggtga tgattgtgtg gcaagtagac
4441 aggatgagga ttagaacatg gaaaagttta gtaaaacacc atatgtatgt ttcagggaaa
4501 gctaggggat ggttttatag acatcactat gaaagccctc atccaagaat aagttcagaa
4561 gtacacatcc cactagggga tgctagattg gtaataacaa catattgggg tctgcataca
4621 ggagaaagag actggcattt gggtcaggga gtctccatag aatggaggaa aaagagatat
4681 agcacacaag tagaccctga actagcagac caactaattc atctgtatta ctttgactgt
4741 ttttcagact ctgctataag aaaggcctta ttaggacaca tagttagccc taggtgtgaa
4801 tatcaagcag gacataacaa ggtaggatct ctacaatact tggcactagc agcattaata
4861 acaccaaaaa agataaagcc acctttgcct agtgttacga aactgacaga ggatagatgg
4921 aacaagcccc agaagaccaa gggccacaga gggagccaca caatgaatgg acactagagc
4981 ttttagagga gcttaagaat gaagctgtta gacattttcc taggatttgg ctccatggct
5041 tagggcaaca tatctatgaa acttatgggg atacttgggc aggagtggaa gccataataa
5101 gaattctgca acaactgctg tttatccatt ttcagaattg ggtgtcgaca tagcagaata
5161 ggcgttactc gacagaggag agcaagaaat ggagccagta gatcctagac tagagccctg
5221 gaagcatcca ggaagtcagc ctaaaactgc ttgtaccaat tgctattgta aaaagtgttg
5281 ctttcattgc caagtttgtt tcataacaaa agccttaggc atctcctatg gcaggaagaa
5341 gcggagacag cgacgaagac ctcctcaagg cagtcagact catcaagttt ctctatcaaa
5401 gcagtaagta gtacatgtaa tgcaacctat acaaatagca atagtagcat tagtagtagc
5461 aataataata gcaatagttg tgtggtccat agtaatcata gaatatagga aaatattaag
5521 acaaagaaaa atagacaggt taattgatag actaatagaa agagcagaag acagtggcaa
5581 tgagagtgaa ggagaaatat cagcacttgt ggagatgggg gtggagatgg ggcaccatgc
5641 tccttgggat gttgatgatc tgtagtgcta cagaaaaatt gtgggtcaca gtctattatg
5701 gggtacctgt gtggaaggaa gcaaccacca ctctattttg tgcatcagat gctaaagcat
5761 atgatacaga ggtacataat gtttgggcca cacatgcctg tgtacccaca gaccccaacc
5821 cacaagaagt agtattggta aatgtgacag aaaattttaa catgtggaaa aatgacatgg
5881 tagaacagat gcatgaggat ataatcagtt tatgggatca aagcctaaag ccatgtgtaa
5941 aattaacccc actctgtgtt agtttaaagt gcactgattt gaagaatgat actaatacca
6001 atagtagtag cgggagaatg ataatggaga aaggagagat aaaaaactgc tctttcaata
6061 tcagcacaag cataagaggt aaggtgcaga aagaatatgc atttttttat aaacttgata
6121 taataccaat agataatgat actaccagct atacgttgac aagttgtaac acctcagtca
6181 ttacacaggc ctgtccaaag gtatcctttg agccaattcc catacattat tgtgccccgg
6241 ctggttttgc gattctaaaa tgtaataata agacgttcaa tggaacagga ccatgtacaa
6301 atgtcagcac agtacaatgt acacatggaa ttaggccagt agtatcaact caactgctgt
6361 taaatggcag tctggcagaa gaagaggtag taattagatc tgccaatttc acagacaatg
6421 ctaaaaccat aatagtacag ctgaaccaat ctgtagaaat taattgtaca agacccaaca
6481 acaatacaag aaaaagtatc cgtatccaga gaggaccagg gagagcattt gttacaatag
6541 gaaaaatagg aaatatgaga caagcacatt gtaacattag tagagcaaaa tggaataaca
6601 ctttaaaaca gatagatagc aaattaagag aacaatttgg aaataataaa acaataatct
6661 ttaagcagtc ctcaggaggg gacccagaaa ttgtaacgca cagttttaat tgtggagggg
6721 aatttttcta ctgtaattca acacaactgt ttaatagtac ttggtttaat agtacttgga
6781 gtactaaagg gtcaaataac actgaaggaa gtgacacaat caccctccca tgcagaataa
6841 aacaaattat aaacatgtgg caggaagtag gaaaagcaat gtatgccccc cccatcagtg
6901 gacaaattag atgttcatca aatattacag ggctgctatt aacaagagat ggtggtaata
6961 gcaacaatga gtccgagatc ttcagacctg gaggaggaga tatgagggac aattggagaa
7021 gtgaattata taaatataaa gtagtaaaaa ttgaaccatt aggagtagca cccaccaagg
7081 caaagagaag agtggtgcag agagaaaaaa gagcagtggg aataggagct ttgttccttg
7141 ggttcttggg agcagcagga agcactatgg gcgcagcgtc aatgacgctg acggtacagg
7201 ccagacaatt attgtctggt atagtgcagc agcagaacaa tttgctgagg gctattgagg
7261 cgcaacagca tctgttgcaa ctcacagtct ggggcatcaa gcagctccag gcaagaatcc
7321 tggctgtgga aagataccta aaggatcaac agctcctggg gatttggggt tgctctggaa
7381 aactcatttg caccactgct gtgccttgga atgctagttg gagtaataaa tctctggaac
7441 agatttggaa taacatgacc tggatggagt gggacagaga aattaacaat tacacaagct
7501 taatacactc cttaattgaa gaatcgcaaa accagcaaga aaagaatgaa caagaattat
7561 tggaattaga taaatgggca agtttgtgga attggtttaa cataacaaat tggctgtggt
7621 atataaaatt attcataatg atagtaggag gcttggtagg tttaagaata gtttttgctg
7681 tactttctgt agtgaataga gttaggcagg gatattcacc attatcgttt cagacccacc
```

```
     7741 tcccaatccc gaggggaccc gacaggcccg aaggaataga agaagaaggt ggagagagag
     7801 acagagacag atccattcga ttagtgaacg gatccttagc acttatctgg gacgatctgc
     7861 ggagcctgtg cctcttcagc taccaccgct tgagagactt actcttgatt gtaacgagga
     7921 ttgtggaact tctgggacgc agggggtggg aagccctcaa atattggtgg aatctcctac
     7981 agtattggag tcaggagcta aagaatagtg ctgttagctt gctcaatgcc acagctatag
     8041 cagtagctga ggggacagat agggttatag aagtagtaca aggagcttat agagctattc
     8101 gccacatacc tagaagaata agacagggct tggaaaggat tttgctataa gatgggtggc
     8161 aagtggtcaa aaagtagtgt ggttggatgg cctgctgtaa gggaaagaat gagacgagct
     8221 gagccagcag cagatggggt gggagcagca tctcgagacc tagaaaaaca tggagcaatc
     8281 acaagtagca acacagcagc taacaatgct gattgtgcct ggctagaagc acaagaggag
     8341 gaggaggtgg gttttccagt cacacctcag gtacctttaa gaccaatgac ttacaaggca
     8401 gctgtagatc ttagccactt tttaaaagaa aaggggggac tggaagggct aattcactcc
     8461 caacgaagac aagatatcct tgatctgtgg atctaccaca cacaaggcta cttccctgat
     8521 tagcagaact acacaccagg gccagggatc agatatccac tgacctttgg atggtgctac
     8581 aagctagtac cagttgagcc agagaagtta gaagaagcca acaaaggaga gaacaccagc
     8641 ttgttacacc ctgtgagcct gcatggaatg gatgacccgg agagagaagt gttagagtgg
     8701 aggtttgaca gccgcctagc atttcatcac atggcccgag agctgcatcc ggagtacttc
     8761 aagaactgct gacatcgagc ttgctacaag ggactttccg ctggggactt tccagggagg
     8821 cgtggcctgg gcgggactgg ggagtggcga gccctcagat cctgcatata agcagctgct
     8881 ttttgcctgt actgggtctc tctggttaga ccagatctga gcctgggagc tc
//
```

NCBI

Sequence Revision History

Entrez

Search for Genes
Entrez Gene provides gene-specific data for multiple taxa

Help|FAQ

Batch Entrez: Upload a file of GI or accession numbers to retrieve protein or nucleotide sequences

Check sequence revision history

How to create WWW links to Entrez

LinkOut

My NCBI (Cubby)

Related resources

BLAST

Reference sequence project

Entrez Gene

Clusters of orthologous groups

Protein reviews on the web

Revision history for M15654

GI	Version	Update Date	Status	I	II
326383	1	Aug 2 2006 2:51 PM	Live	◉	○
326383	1	Oct 4 1994 9:13 AM	Dead	○	◉
326383	1	Aug 2 1993 8:33 PM	Dead	○	○

Accession M15654 was first seen at NCBI on Aug 2 1993 8:33 PM